MULTIMEDIA UNIVERSITY STUDENT IDENTIFICATION NO

MULTIMEDIA UNIVERSITY

FINAL EXAMINATION

TRIMESTER 2, 2015/2016

BST 2044 – APPLIED BUSINESS STATISTICS

(All sections / Groups)

3 MARCH 2016
9.00 AM – 11.00 AM
(2 Hours)

INSTRUCTIONS TO STUDENT

1. This question paper consists of 8 pages with 4 questions only.
2. Attempt **ALL** the questions. All questions carry equal marks and the distribution of the marks for each question is given.
3. Students are allowed to use scientific calculators (without restriction).
4. Statistical formulae are attached at the end of the question paper.
5. Please write all your answers in the Answer Booklet provided.

QUESTION 1

(a) A branch of a bank located in a commercial district of a city has the business objective of improving the process for serving customers during the 12pm to 1pm lunch period. The waiting time (defined as the time the customer enters the line until he reaches the teller counter) of a random sample of 24 customers is collected. The results are shown as below:

4.21	5.67	2.34	6.08	4.76	2.89	4.45	3.89
3.34	4.36	3.55	2.67	4.41	4.22	2.04	3.49
4.56	3.47	3.91	5.32	5.51	3.23	4.11	3.66

(i) Construct a 90% confidence interval estimate for the mean waiting time of the customer during the lunch period. (7 marks)

(ii) Is there any significant evidence that the average waiting time is less than 4.5 minutes at the 0.01 significance level? (8 marks)

(b) A PC company uses two suppliers for notebook rechargeable batteries. Two important factors of the quality features of the batteries are mean use time and variation. It is desirable that the mean use time to be high and the variability to be low. Recently, the PC maker conducted a test on batteries from these suppliers. 9 randomly selected batteries from Supplier 1 and 12 randomly selected batteries from Supplier 2 were tested. The following results were observed:

Supplier 1	**Supplier 2**
$n_1 = 9$	$n_2 = 12$
$\bar{x}_1 = 67.25$ min.	$\bar{x}_2 = 72.4$ min.
$s_1 = 11.2$ min	$s_2 = 9.9$ min.

Based on these sample results, can the PC maker conclude that a difference exists between the two batteries with respect to the population standard deviations? Test using $\alpha = 0.10$.

(10 marks)

(Total: 25 marks)

Continued...

QUESTION 2

(a) An investment advisor is interested to know whether there is a difference in the annualised average returns (in RM) between four different investment portfolios: Aggressive, Balance, Growth and Conservative. The summary and analysis of variance (ANOVA) of the portfolios are shown in the following:

SUMMARY

Portfolio	*No. of stocks*	*Average return*	*Variance*
Aggressive	16	431.50	150392.27
Balance	16	407.31	61055.96
Growth	15	421.93	63002.92
Conservative	12	408.25	54493.66

ANOVA

Source of Variation	*SS*	*df*	*MS*
Between Groups	***	3	***
Within Groups	***	***	***
Total	4659308.1356	58	

(i) Compute the Sum of Squares Within (*SSW*) the portfolios and hence complete the ANOVA table. (6 marks)

(ii) By assuming that the populations have equal variances and are normally distributed, what conclusion can be made by the investment advisor at the 5% significance level? (7 marks)

(b) An economist would like to examine whether there is a relationship between Melaka's foreign direct investments (rounded to the nearest million) and its level of crime rates (High, Moderate, and Low). The related information of fifty randomly selected foreign investors is provided in the following table:

Level of crime rates	Foreign direct investments			
	1 – 10	11 – 20	21 – 30	> 30
High	3	2	4	1
Moderate	6	4	3	2
Low	7	8	6	4

Is there a relationship between Melaka's foreign direct investments and its level of crime rates at the 10% significance level? (12 marks)

(Total: 25 Marks)

Continued...

QUESTION 3

(a) The marketing manager of a Hot Dog Stand chain is in the process of examining some of the demographic characteristics of her customers. In particular, she would like to investigate the belief that the ages of the customers of hot dogs, fast-food hamburger, and fast-food chicken restaurants are different. To test this belief, the ages of six randomly selected customers from each restaurant type were recorded and listed below. From previous analysis we know that the ages are not normally distributed.

Customers' Ages

Hot Dogs	Hamburger	Chicken
23	26	36
19	20	23
17	18	39
36	35	27
25	33	38
28	19	31

Does this data provide enough evidence at the 10% significance level to infer that there are differences in ages among the customers of the three restaurant types?

(13 marks)

(b) A real estate broker claims that the median days of one of his listings stays on the market is 45 or less days. To test this, he has collected the following random sample of properties sold showing the days they were on the market prior to selling:

50	30	70	20	30	40	60	80

The broker is unwilling to assume that the population data are normally distributed Use Wilcoxon signed-rank test in evaluating the broker's claim at the 0.05 level of significance.

(12 marks)

(Total: 25 marks)

Continued...

QUESTION 4

(a) The manager of a fast food restaurant would like develop a model to predict the time it takes to deliver food to their customer. Data were collected from the past 15 orders, the delivery time and distance from the restaurant are as following

Distance (KM)	1	2.7	9.6	6.2	6.7	9.3	2	1.6	3.9	8.9	1.7	2.4	7.8	3.9	9.3
Delivery Time (Minutes)	3	8	27	14	16	26	6	5	10	22	5	7	19	11	25

(i) Compute the coefficient of correlation and interpret your answer. (4 marks)

(ii) Find the regression line of the delivery time. (5 marks)

(iii) Estimate the time needed to deliver food to a customer who is 6km away from the restaurant. Is this estimation reliable? (2 marks)

(b) The following ANOVA table is obtained from a multiple linear regression model with five independent variables:

ANOVA Table

Source of Variation	***df***	***Sum of Squares***	***Mean Squares***
Regression	**	**	**
Residual	**	120.22	**
Total	24	250.32	

(i) Complete the ANOVA table. (2 marks)

(ii) Test whether the regression model is statistically significant at 5% significance level. (5 marks)

(c) The numbers of passengers arriving at Melaka International Airport on 6 successive Mondays were: 60, 72, 96, 84, 36, and 48.

Construct an exponential smoothing model for the number of arrivals above data by using a smoothing constant of 0.25.

(7 marks)

(Total: 25 marks)

End of Page

STATISTICAL FORMULAE

Descriptive Statistics

Mean = $\frac{\sum x}{n}$

Standard Deviation (s) = $\sqrt{\frac{\sum x^2}{n-1} - \frac{(\sum x)^2}{n(n-1)}}$

Confidence Interval Estimation and Sample Size Determination

(100 - α)% Confidence Interval for Population Mean,(σ Known) is

$\mu = \bar{x} \pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$

(100 - α)% Confidence Interval for Population Mean (σ Unknown) is

$\mu = \bar{x} \pm t_{\alpha/2, n-1} \frac{s}{\sqrt{n}}$

(100 - α)% Confidence Interval for Population Proportion is

$\pi = p \pm Z_{\alpha/2} \sqrt{\frac{p(1-p)}{n}}$

Finite Population Correction Factor = $\sqrt{\frac{N-n}{N-1}}$

Sample Size Determination for Population Mean = $n = \left(\frac{Z_{\alpha/2}\sigma}{e}\right)^2$

Sample Size Determination for Population Proportion = $n = \frac{(Z_{\alpha/2})^2 p(1-p)}{e^2}$

Where e = Limit of Error in Estimation

Hypothesis testing

One Sample Mean Test	
(Standard Deviation (σ) Known)	**(Standard Deviation (σ) Not Known)**
$Z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$	$t = \frac{\bar{x} - \mu}{s / \sqrt{n}}$
One Sample Proportion Test	
$Z = \frac{p - \pi}{\sqrt{\frac{\pi(1-\pi)}{n}}}$	

Test for the equality of Two Variances

$$F = \frac{s_1^2}{s_2^2}$$

Chi Square Test

$$\chi^2 = \sum \frac{(o_i - e_i)^2}{e_i}$$

Wilcoxson Rank Sum Test

$$Z_{T_1} = \frac{T_1 - \mu_{T_1}}{\sigma_{T_1}} \text{ Where } \mu_{T_1} = \frac{n_1(n_1 + n_2 + 1)}{2} \text{ and } \sigma_{T_1} = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}$$

Kruskal Wallis Test

$$H = \left[\frac{12}{n(n+1)} \sum_{i=1}^{k} \frac{R_j^2}{n_i} \right] - 3(n+1) \text{ and } H \sim \chi^2_{k-1}$$

Wilcoxson Signed Rank Sum Test

$$Z_{T_{+1}} = \frac{T_+ - \mu_{T_{+1}}}{\sigma_{T_+}} \text{ Where } \mu_{T_{+1}} = \frac{n(n+1)}{4} \text{ and } \sigma_{T_+} = \sqrt{\frac{n(n+1)(2n+1)}{24}}$$

Sign Test

$$Z = \frac{(n_+ - 0.5n)}{\sqrt{n(0.5)(0.5)}}$$

C. One –Way ANOVA

Source	Sum of Squares	Degree of Freedom	Mean Square	F-statistic
Among Groups	SSA	k-1	MSA=SSA/k-1	MSA/MSW
Within Groups	SSW	N-k	MSW=SSW/N-k	
Total	SST	N-1		

SST = SSA + SSW $\qquad$ $SST = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (x_{ij} - \bar{x})^2$

$SSA = \sum_{i=1}^{k} n_i (\bar{x}_i - \bar{\bar{x}})^2$ $\qquad$ $SSW = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2$ or $SSW = \sum_{i=1}^{k} (n_i - 1) s_i^2$

n = Number of observations in each group and N = Total number of observations

D. Regression Analysis

Simple Linear Regression

$$Y = a + bX, \quad \text{where } b = \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sum x^2 - \frac{(\sum x)^2}{n}} \text{ and } a = \bar{y} - b\bar{x}$$

Correlation coefficient

$$r = \frac{n\sum xy - \sum x \sum y}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}}$$

Adjusted r-square =1 - [(1- r^2) (n-1) / (n-p-1)] Where p =number of independent variables

ANOVA Table for Regression

Source	**Degree of Freedom**	**Sum of Squares**	**Mean Squares**
Regression	K	SSR	MSR=SSR/k
Residual	n-k-1	SSE	MSE=SSE/(n-k-1)
Total	n-1	SST	

Test Statistic for Significance of the Overall Regression Model = F= MSR/MSE

Test Statistic for Significance of each Explanatory Variable = **t** = (β_i - 0) / $S_{\beta i}$ and the Critical t = $t_{(n-P-1),\alpha/2}$

Wilcoxon Matched-Pairs Signed-Ranks Test Critical Values T ($n \leq 25$)

n	LEVEL OF SIGNIFICANCE FOR ONE-TAILED TEST		
	0.025	0.01	0.005
	LEVEL OF SIGNIFICANCE FOR TWO-TAILED TEST		
	0.05	0.02	0.01
6	0	—	—
7	2	0	—
8	4	2	0
9	6	3	2
10	8	5	3
11	11	7	5
12	14	10	7
13	17	13	10
14	21	16	13
15	25	20	16
16	30	24	20
17	35	28	23
18	40	33	28
19	46	38	32
20	52	43	38
21	59	49	43
22	66	56	49
23	73	62	55
24	81	69	61
25	89	77	68

Wilcoxon Signed-Ranks Test Lower and Upper Critical Values W

	One-Tailed: $\alpha = .05$	$\alpha = .025$	$\alpha = .01$	$\alpha = .005$
	Two-Tailed: $\alpha = .10$	$\alpha = .05$	$\alpha = .02$	$\alpha = .01$
n	(Lower, Upper)			
5	0,15	—,—	—,—	—,—
6	2,19	0,21	—,—	—,—
7	3,25	2,26	0,28	—,—
8	5,31	3,33	1,35	0,36
9	8,37	5,40	3,42	1,44
10	10,45	8,47	5,50	3,52
11	13,53	10,56	7,59	5,61
12	17,61	13,65	10,68	7,71
13	21,70	17,74	12,79	10,81
14	25,80	21,84	16,89	13,92
15	30,90	25,95	19,101	16,104
16	35,101	29,107	23,113	19,117
17	41,112	34,119	27,126	23,130
18	47,124	40,131	32,139	27,144
19	53,137	46,144	37,153	32,158
20	60,150	52,158	43,167	37,173